Journal of Memory and Language 135 (2024) 104478

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Language comprehenders are sensitive to multiple states of semantically similar objects

Oleksandr V. Horchak *, Margarida V. Garrido

Iscte - Instituto Universitário de Lisboa, Cis-Iscte, Portugal

ARTICLE INFO

Keywords:
Language comprehension
Event representation
Object state
Semantic similarity

ABSTRACT

The present research shows that language comprehenders are sensitive to multiple states of target and semantically related objects. In Experiments 1 to 2B, participants (total $N = 273$) read sentences that either implied a minimal change of an object's state (e.g., "Jane *chose* a mango") or a substantial change (e.g., "Jane *stepped* on a mango") and then verified whether a subsequently pictured object was mentioned in the sentence. Crucially, the picture either showed the original/modified state of an object that was mentioned in the sentence (e.g., "mango" in Experiment 1) or not (e.g., "banana" in Experiments 2A and 2B). The results of Experiment 1 demonstrated that the objects in a modified state were verified faster when a sentence implied a substantial state-change rather than a minimal state-change. In contrast, the reverse was true for the objects in the original state. Importantly, verification latencies of pictures depicting original and modified states of an object in the substantial state-change condition were approximately similar, thus suggesting that language comprehenders maintain multiple representations of an object in different states. The results of Experiments 2A and 2B revealed that when participants had to indicate that a pictured object (e.g., banana) was not mentioned in the sentence, their verification latencies were slowed down when the sentence contained a semantically related item (e.g., mango) and described this item as being changed substantially by the action. However, these verification latencies varied continuously with the degree of change: the more dissimilar the states of a semantically related item, the less time participants needed to verify a pictured object. The results are discussed through the prism of theories emphasizing dynamic views of event cognition.

Introduction

Many theories of event cognition and language processing suggest that comprehending an event is tantamount to constructing a mental representation of the described situation (Johnson-Laird, 1983; van Dijk & Kintsch, 1983; Zwaan, Langston, & Graesser, 1995; Zwaan & Radvansky, 1998). According to these theories, language comprehenders place themselves in the described situation and experience the event on a number of dimensions, including those related to protagonists (e.g., Gernsbacher et al., 1992), space (e.g., Glenberg et al., 1987), time (e.g., Zwaan, 1996), and goals (e.g., Trabasso & Suh, 1993). For example, there is now quite a body of data from a sentence-picture verification paradigm showing that people rapidly integrate implied visual information about the object's orientation (e.g., a nail in the wall vs. a nail in the floor) and shape (e.g., a bird in the nest vs. a bird in the sky; e.g., Stanfield & Zwaan, 2001; Zwaan et al., 2002). Furthermore, there are studies demonstrating that people also simulate a more global environmental context, such as the setting in which an event takes place (e.g., Horchak & Garrido, 2022; Horton & Rapp, 2003; Yaxley & Zwaan, 2007). Altogether, these data converge on the conclusion that comprehenders draw information from the surrounding environment to be able to successfully understand language.

However, surrounding context (e.g., location, background settings, etc.) is not the only significant factor. Another less-studied but equally fundamental element of event comprehension is object history, as suggested by the Intersecting Object Histories (IOH) account (Altmann & Ekves, 2019). According to this account, comprehenders need to encode changes in object state (i.e., the association of objects with their past selves through space and time) to build a rich representation of an event. Evidence for this can be found in behavioral studies showing match effects between visual and linguistic information when implied object state-changes are compatible rather than incompatible. For instance, Kang et al. (2020) used a sentence-picture verification task to test the hypothesis that verbs may drive the updating of state information during

* Corresponding author at: Iscte - University Institute of Lisbon, Av. das Forças Armadas, 1649-026 Lisbon, Portugal.
E-mail address: Oleksandr.Horchak@iscte-iul.pt (O.V. Horchak).

https://doi.org/10.1016/j.jml.2023.104478
Received 15 March 2023; Received in revised form 6 November 2023; Accepted 14 November 2023

sentence comprehension. Specifically, they asked participants to read sentences like "Jane *chose* a banana" or "Jane *stepped* on a banana" and then indicate whether the subsequently presented pictured object was mentioned in the sentence. By segregating the data by pictures, they found that response times (RT) were faster whenever there was a match between the final state of an object depicted in the picture and the one implied by the past-tense sentence. Interestingly, however, Kang et al. (2020) did not find the match advantage for the original object state when future-tense sentences (e.g., "Jane *will choose* a banana" vs. "Jane *will step* on a banana") were used, thus suggesting that the associated representation of the initial (original) state was more accessible than the future (modified) state (see also Lee & Kaiser, 2021; Misersky et al., 2021, for related findings on how language comprehenders focus more on the initial object state when a sentence is in future tense). Hence, the results are consistent with the notion of affordances, which refers to the possibility of an action on an object (Gibson, 1979; Symes et al., 2007). Specifically, future-tense sentences suggest equal affordance for both substantial-change (i.e., "will *step* on a banana") and minimal-change ("will choose a banana") conditions precisely because the subject intends to act on the object (i.e., "banana") in its original state from the subject-centric current (not future) state of the world. Most importantly, Kang et al.'s (2019) results not only showed that appropriate object representations are modulated by the degree of implied state-change but also that multiple representations of the same object in the different states are simultaneously active. The evidence for this is found when looking at the data segregated by sentences. Specifically, the authors demonstrated that when the object was described as being changed substantially by the action (e.g., "Jane *stepped* on a banana"), participants' verification latencies of the objects in both the original (i.e., intact banana) and the modified states (i.e., smashed banana) were approximately the same. Furthermore, no differences in RTs in the substantial-change condition were also reported in a related study by Horchak and Garrido (2021) on the effect of the object state being affected by light vs. heavy items (e.g., "He dropped a *bowling ball* on a banana" vs. "He dropped a *balloon* on a banana"). Finally, neuroscientific data revealed that a neural marker for competition is present more when participants process the sentences in the substantial-change condition than in the minimal-change condition (Hindy et al., 2012; Solomon et al., 2015). Thus, these data suggest that comprehension of an action in a sentence like "The woman *stepped* on a banana" requires activation at the "banana" of both the original (canonical) state of a banana and of the modified (non-canonical) state of a banana – the consequence of a woman smashing it. Accordingly, maintaining multiple representations of an object in different states may engender a conflict precisely because it requires the comprehenders to choose the situationally appropriate object state among the other states through which it has passed.

Thus far, the available evidence offers empirical support for the claim that the processing of object state-change denoted by the verb (e.g., step) requires activating some representation of the transition from one state of an object (e.g., banana) to another. There are limitations, however, to what we know about the role of multiple object-state representations during language processing. Specifically, it is currently unclear whether the activation of these multiple object-state representations may have some consequences for the processing of semantically related (as compared to unrelated) concepts. This issue is important to address as, in most cases, there is more than just one participant (object) in the event.

We follow Altmann and Ekves (2019) in assuming that the representation of an object's history is grounded in episodic experiences through the process of relational binding. The central tenet of this process is that we experience objects in the presence of other objects and background settings. Therefore, our experience of such encounters should include incidental properties of the encounter, as well as other elements with which they are associated in semantic memory. So, on reading sentences such as "Jane *stepped* on a mango" and "Jane *chose* a mango", such incidental properties may include the places where this fruit can be found (tropical environments), the culinary dishes in which this fruit can be used (desserts), other fruits associated with mangoes (e.g., bananas), and the different states through which these different fruits have (could have) passed (e.g., slicing or blending). Consequently, if language comprehenders bind activated semantic representations (e.g., reading "mango" activates semantic knowledge of food) to the episodic contexts that lead to their activation (e.g., kitchen setting), as well as to other elements with which they are associated in semantic memory (e.g., mangoes are associated with bananas, papayas, etc.), then they could be sensitive not only to the interference caused by similar items (e.g., a mango vs. a banana) but also to the interference caused by different states of the same item (e.g., intact mango vs. squashed mango). The present study takes a first step towards addressing this prediction. Specifically, we investigated whether language comprehenders are sensitive to the linguistically determined states of an object (e.g., choose vs. step on a mango) when verifying a semantically similar pictured object (e.g., banana in either original or modified state).

Before turning to our experiments, we will first outline previous empirical evidence that served as the basis for our predictions. First, we will review prior research on thematic and taxonomic relatedness and clarify our definition of semantic similarity in the present paper. Second, we will turn to studies focusing on the resolution of state ambiguity (e.g., representing the banana *before* or *after* it was stepped on) during language comprehension. Specifically, our focus will be on studies demonstrating how representing information from a different, yet related, object may be difficult due to similarity-based interference (i.e., interference caused by the difference between distinct objects) and/or dissimilarity-based interference (i.e., interference caused by the '*before*' and '*after*' object states).

Defining semantic similarity

Before moving forward, it is important to clarify our treatment of semantic similarity in the present paper. We shall do so by highlighting the distinctions between two major kinds of semantic relations: taxonomic and thematic. Taxonomic similarity refers to the relationship between words or concepts that belong to the same hierarchical category, whereby similarity between concepts is defined as a function of feature overlap (e.g., Cree et al., 1999; Mervis & Rosch, 1981; Smith et al., 1974). Thematic similarity, on the other hand, is based on grouping concepts thematically based on temporal, spatial, causal, or functional relations between things (e.g., Golonka and Estes, 2009, Lin and Murphy, 2001). In contrast to taxonomic similarity, objects that share thematic relations, such as coffee and newspaper (morning routine) or pen and paper (writing notes), often share a few (if any) features. Instead, they have complementary attributes that correspond to their distinct roles within specific events or situations.

Substantial empirical evidence shows that both taxonomic and thematic relations are essential to language comprehension and conceptual processing (see Estes et al., 2011; Mirman et al., 2017, for comprehensive reviews). A common trend that emerges from various studies using picture stimuli, for example, is that participants are slowed down when the target picture is accompanied by a semantically similar rather than unrelated distractor. In language production literature, this result is usually interpreted as evidence for a competitive lexical selection process (see Nozari & Pinet, 2020, for discussion). Similarly, in the language processing literature, this finding is claimed as evidence for semantic similarity between conceptual knowledge accessed by the distractor and the conceptual knowledge accessed from the target object (e.g., Huettig & Altmann, 2005). However, another important trend found in the literature is that taxonomic (feature-based) and thematic (associative) similarities do not always have the same effect on behavior. Some research on language processing shows both taxonomic and thematic competition effects, with the former being larger (e.g., Mirman & Graziano, 2012). At the same time, other research on language production, for example, shows that distractors that are thematically (not taxonomically) associated with a target may, in fact, lead to facilitation (e.g., Mahon et al., 2007). Thus, the precise nature of the semantic relationship between the target and the distractor (i.e., taxonomic or thematic relationship) could, at the very least, affect the size of the observed competitor effect.

Many studies on language comprehension focus on thematic relations that help us generate expectations about events or scenarios (e.g., Cooper, 1974; Huettig & Altmann, 2005). As nicely pointed out by Estes et al. (2011), while knowing that a menu is taxonomically related to a book is useful (e.g., both have pages), it is equally important to know how to guide your behavior with respect to the event. In other words, thematically related items such as food, waiters, water, or wine may share few features with a menu but are nonetheless important to generate expectations within an event. Corpus-based semantic distance measures like, for example, Latent Semantic Analysis (LSA) were found to be particularly well-suited for determining thematic relationships between words. In straightforward terms, LSA (Landauer & Dumais, 1997) takes many words from different documents, finds hidden patterns, and mathematically estimates how similar words are based on where they are placed in this pattern space. Ultimately, the similarity between the words is derived from the pattern of frequencies of pairs of words across different documents: words that are semantically related have higher LSA scores than words that are unrelated. As one illustrative example, Huettig et al. (2006) demonstrated that LSA scores predicted eye fixation behavior in the visual world paradigm. Specifically, they first calculated the similarity between various words using the LSA method. Then, they tested how the degree of similarity between two words (target and competitor) affects fixation behavior. As expected, the authors found that on hearing a sentence containing a critical target word (e.g., toaster), participants showed a bias to fixate a picture of a semantically related (as defined by LSA) competitor item (e.g., corkscrew) in proportion to their LSA similarity.

Given the above evidence on the reliability of semantic distance models, in the present research, we used a corpus-based word embedding technique, *word2vec,* to determine semantic similarity (e.g., Mikolov et al., 2013). Although this embedding technique doesn't directly differentiate between taxonomic and thematic relationships (i.e., *word2vec* embeddings tend to cluster words that are taxonomically similar or thematically related together in the vector space), it is widely used to predict lexical associations (i.e., the connections or relationships between words) based on the co-occurrence of words in similar contexts that share common themes or topics. Thus, our treatment of semantic similarity in the present paper is consistent with corpus-based semantic distance measures where thematic similarity plays an important role.

State ambiguity and similarity-based interference

Why should language comprehenders be sensitive to states of semantically similar objects? Our starting point is the above empirical evidence that there is an overlap between the conceptual information conveyed by words and the conceptual knowledge associated with semantically similar (pictured) objects. Consider, for example, a situation presented in Fig. 1 when the participants' task is to read the sentences describing different states of an object (e.g., bulb vs. mango) and then judge whether the subsequently pictured item (e.g., a banana) in either the original or modified state had been mentioned in the sentence. All sentences describe similar situations except for one fundamental difference: the pictured banana has a stronger semantic relationship with the word "mango" than with the word "bulb". In line with the evidence that comprehenders are sensitive to semantically similar objects (e.g., McRae & Boisvert, 1998; Yee & Thompson-Schill, 2016), participants' verification latencies of a banana should be slower after reading a sentence mentioning a mango rather than a bulb, precisely due to difficulties in resolving a competition between shared characteristics (e.g., both are fruits, both grow in tropical climates, both can be used in desserts, etc.) of the target (i.e., a banana) and a semantically similar (i.

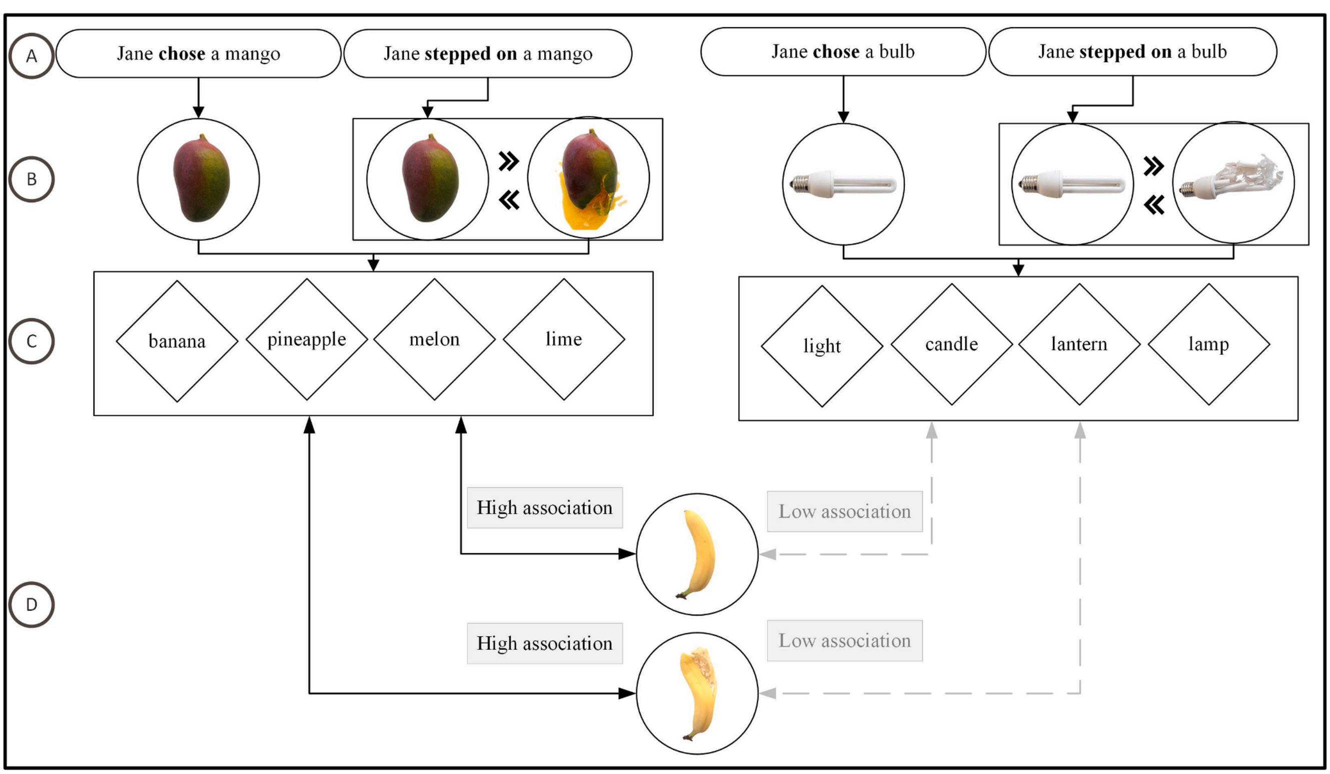

Fig. 1. An Example of Associations that Form When Participants Perform a Task Note. Arrows show the interplay between perceptual properties of objects and general semantic knowledge about objects. **A:** Examples of sentences with different verb information. **B:** Required representations of object states during event comprehension. **C:** Example of semantic knowledge being activated during event comprehension. **D:** Examples of pictured objects in different states (original vs. modified) that are either highly or weakly associated with the object from the sentence.

e., a mango) concept. In contrast, no such difficulties are expected when participants read sentences such as "Jane *stepped* on a bulb" and "Jane *chose* a bulb", given that the conceptual representations activated by the word "bulb" only minimally overlap with those associated with the target object "banana". Thus, on this view, greater semantic similarity between the items leads to greater *similarity-based* interference (see Damian & Bowers, 2003; Schriefers et al., 1990, for further discussion of semantic interference).

However, according to the view advocated in the present paper, participants' verification latencies of the picture of a banana should be slowed down more after reading a sentence where a semantically related object undergoes a change of state than after reading a sentence where a semantically related object does not undergo a change of state. This is expected in light of the amount of semantic priming afforded by a sentence describing a substantial change in an object's state. More specifically, and in line with the IOH account (Altmann & Ekves, 2019), we propose that after reading a sentence like "Jane *stepped* on a mango", multiple representations of the same object (i.e., intact and squashed mango), reflecting the different states it passes through, are activated. In contrast, after reading the sentence, "Jane *chose* a mango", only one (i.e., intact) object-state representation is activated. As a result, a semantically related banana gets more priming from the "intact and squashed mango" than just from the "intact mango", and this leads to longer RTs. To put it differently, it should take longer to reject the banana after a sentence referring to a mango than to one referring to a bulb because the feature set of the banana, overlapping as it does with the feature set of the mango, is harder to distinguish from that of the mango.

Providing evidence for the amount of semantic priming afforded by the "substantial state-change" sentence condition is theoretically important: if linguistically determined atypical exemplars (e.g., the "step on" sentence where the object is implied to be in the less typical modified state) may be better primes than typical exemplars (e.g., the "choose" sentence where the object is implied to be in the more typical intact state), then this would run counter to the views that typicality always boosts semantic priming effects (see, for example, Collins & Loftus, 1975; Rosch, 1975; for assumptions of spreading activation theory and the prototype model, respectively). Also, this would counter the proposal that concepts are represented in terms of fixed points in semantic space (e.g., Landauer & Dumais, 1997). Rather, this would lend further credence to the empirical evidence emphasizing the non-static nature of semantic cognition, which showed, among other things, that language context provides a strong constraint on the activation of conceptual properties (e.g., Rogers & McClelland, 2005; Smith et al., 1974; Huettig & Altmann, 2005; Mirman & Magnuson, 2009; Mirković & Altmann, 2019; Yee & Sedivy, 2006).

State ambiguity and Dissimilarity-Based interference

Whereas similarity-based interference may explain how language comprehenders resolve state ambiguity of distinct objects as a function of different action verbs (e.g., choose vs. step on), it may not be able to fully explain how they resolve state ambiguity depending on the extent to which any specific object changes in state. For example, both stepping on a mango and stepping on a papaya suggest a substantial change of state. However, the consequences of stepping on these two fruits will likely differ due to their contrasting physical properties. Mangoes are relatively firmer compared to papayas. When stepped on, a mango is likely to offer more resistance and might not flatten as easily. In contrast, papayas have a soft, almost butter-like texture when ripe. Stepping on a papaya would likely result in more immediate deformation and flattening than stepping on a mango. Thus, the representations of intact and squashed papaya will generally be more dissimilar than those of intact and squashed mango.

The relevance of dissimilarity of this kind has been demonstrated using fMRI. Specifically, it was shown that high dissimilarity of object states (i.e., *dissimilarity-based* interference) leads to object rivalry, whereby the representations of the distinct states of the same object interfere with one another during event comprehension. For example, Hindy et al. (2012) provided such evidence by focusing on the left posterior ventrolateral prefrontal cortex (pVLPFC) – the area of the brain that responds to Stroop conflict (e.g., Banich et al., 2000). In the Stroop task (e.g., MacLeod & MacDonald, 2000), participants see a list of color words (e.g., red, blue, green) where the color of the ink used to write the word does not match the actual word. Participants are required to name the ink color while ignoring the written word. This creates a conflict between the automatic processing of reading the word and the controlled processing of identifying the ink color. Thus, Hindy et al. (2012) reasoned that the resolution of state ambiguity should lead to conflict similar to that observed in the Stroop task precisely because both tasks require selecting one representation at the expense of the other.

To test their predictions, Hindy et al. (2012) asked participants to lie in the scanner and read pairs of sentences like (1) "The chef will *chop* the onion. Then she will *weigh* the onion" or (2) "The chef will *smell* the onion. Then she will *weigh* the onion". Given the context, sentences in pair (1) imply different states of an onion (i.e., initial and chopped states), while sentences in pair (2) do not imply any changes in object state. The authors found evidence of conflict when participants read "chop the onion" versus when they read "weigh the onion", given that there was greater left pVLPFC activity during the processing of trials describing a substantial state-change (i.e., the "chop" action). Furthermore, the left pVLPFC response correlated with ratings provided by separate participants of the degree of change that each specific object underwent because of action. Specifically, Hindy et al. (2012) found that the more dissimilar an object's initial and modified states were, the greater the interference.

A related study by Solomon (2015) adopted the experimental paradigm and stimuli of Hindy et al. (2015) to provide more convincing evidence that the representational conflict occurs specifically due to selecting mutually exclusive object states rather than maintaining similar object representations (that is, chopped onion still looks similar to intact onion so long as it is recognizable as an onion, and hence may lead to a similarity-based conflict). The only difference in the materials was the number and type of object referents. Specifically, they presented participants with the sentences in three different conditions: S-token (e.g., "The chef will *chop* the onion. Then she will *weigh* the onion"), D-token ("The chef will *chop* the onion. Then she will *weigh another* onion"), and D-type (e.g., "The chef will *chop* the onion. Then she will *weigh a piece of garlic*"). That is, they contrasted the events involving the same token (S-token condition) with the events involving a different token of the same type (D-token condition) and the events involving two different object types (D-type condition). The major result was that the representational conflict was observed in the S-token condition but not in the D-token condition. This is consistent with the idea that competition of object states arises only when the event describes mutually exclusive states of the same object. That is, when the event describes *another* object (i.e., another onion), as in the D-token condition, maintaining multiple overlapping representations has no inhibitory consequences. Therefore, different objects can coexist no matter how similar they are. Presumably, this reflects the fact that when one is processing information about *another* object (i.e., another onion), there is no need to inherit the episodic characteristics of the first-mentioned object. On the face of it, the results of Solomon et al. (2015) may suggest that dissimilarity-based interference is unlikely to have any implications for the present study with distinct, albeit semantically similar, objects. However, one other finding in the results of Solomon et al. (2015) gives us pause. More precisely, this concerns their finding that left pVLPFC was activated more during the processing of substantial changes than minimal changes for events that include tokens of two different (not the same) concept types (i.e., onion vs. garlic). Solomon et al. (2015) interpreted this finding as evidence of competition between incompatible object types. However, it is currently unclear whether the fMRI evidence for dissimilarity-based interference between object types can also be observed in a behavioral study. The experiments reported in the present paper were designed to provide such initial evidence.

In a functional sense, dissimilarity-based interference is expected to be inhibitory as it requires selecting one state representation at the expense of the other. Consequently, if dissimilarity-based interference is manifested in the degree of dissimilarity between the initial and modified states of an object, then an increase in the RTs to the pictured item (e.g., banana) in the substantial-change condition should scale inversely with the degree of change of the semantically related (e.g., mango) item (as rated by other participants). For example, if the representations of intact and squashed papaya are rated as more dissimilar than the representations of intact and squashed mango, then the object states of the papaya should inhibit one another more than the object states of the mango. Consequently, there should be less priming of the banana after "stepped on a papaya" than after "stepped on a mango", and hence the decreased RTs (i.e., less priming should make it easier for participants to reject the pictured banana).

The present study

In the present study, we tested the hypothesis that language comprehenders are sensitive not only to the interference caused by similar objects (i.e., *similarity-based* interference) but also to the interference caused by different states of the same object (i.e., *dissimilarity-based* interference). Our broader goal was to contribute to a unified theory of event comprehension, where event representations are grounded in representations of the different states of an object. Experiment 1 laid the foundation for Experiments 2A and 2B by targeting the question of whether linguistic context has an impact on the activation of object-state representations. In essence, we first wanted to conceptually replicate the critical behavioral findings of Kang et al. (2020) in support of the IOH account. Specifically, we expected to find that original and modified states of an object would be verified faster when a sentence implies a minimal and a substantial state-change, respectively. Equally important, we wanted to confirm that multiple instances of the same object are activated when participants read a sentence describing a substantial change in object state. Experiments 2A and 2B addressed the main research question regarding the sensitivity of language comprehenders to multiple object state representations of semantically similar objects. Our predictions were as follows: If it is the case that after reading (1) "Jane *stepped* on a mango", multiple representations of the same object are activated, and after reading (2) "Jane *chose* a mango" only one object state is activated, then a semantically related banana, regardless of its pictured state, should get more priming from sentence (1) than from sentence (2), and hence elevated RTs. However, the amount of priming in sentence (1) should also vary continuously with the degree of change (as rated by different participants). Specifically, we expected to observe that the greater the difference between the state of the object implied by a "substantial state-change" sentence (1) and a "minimal state-change" sentence (2), the greater the conflict between two states should be. Consequently, the conflict between highly dissimilar object states (i.e., states that struggle to coexist) should have some inhibitory effect and, as a result, decrease the amount of priming afforded by sentence (1).

To address these questions, we used a sentence-picture verification paradigm where participants were asked to read sentences that either implied a minimal change of an object's state (e.g., "Jane *chose* a banana") or a substantial change (e.g., "Jane *stepped on* a banana") and then verified whether a subsequently pictured object was mentioned in the sentence (i.e., banana). In Experiment 1, participants were expected to provide a" yes" response as the critical pictured object (depicted in either original or modified state) matched the object (either intact or changed) mentioned in the sentence. In Experiments 2A and 2B, participants were expected to provide a "no" response as the critical pictured object mismatched the object (either intact or changed AND either semantically related or unrelated) mentioned in the sentence. The only difference between Experiments 2A and 2B was in the state of the pictured object displayed on the screen: the object was intact in Experiment 2A and modified in Experiment 2B. A detailed overview of the paradigm (in Experiments 1, 2A, and 2B) is provided in Fig. 2.

Experiment 1

The purpose of Experiment 1 was two-fold. First, we wanted to show that the activation of the contextually appropriate object representation is modulated by the degree of state-change denoted by the sentence. Specifically, we expected to find faster verification times for the pictures depicting an object in the original state after sentences describing a minimal change of state; and faster verification times for the pictures depicting an object in the modified state after sentences describing a substantial change of state. Second, we wanted to establish that language comprehenders need to represent multiple instances of the same object in different states. More precisely, we expected to observe similar picture verification latencies for objects in the original (e.g., intact banana) and modified (e.g., smashed banana) states after reading a sentence that describes an object as changing substantially (e.g., "Jane *stepped* on a banana"). Thus, we sought to conceptually replicate the findings of Kang et al. (2019) about the role of verbs as the "driver" of updating state information in sentence comprehension. The major differences in our stimulus materials were as follows. First, in Kang et al. (2019), the sentences started with common nouns (e.g., boy, woman), and in the present study, the sentences started with proper nouns (e.g., Jane, John, etc.). Second, in Kang et al. (2019), the picture stimuli were clip art objects, and in the present research, the picture stimuli were photographs of real objects.

Method

Estimation of sample size

To determine the required sample size, we performed a simulation-based power analysis using the "mixedpower" package of Kumle et al. (2018). Specifically, we used the data from Experiment 7 in Horchak and Garrido (2021), where pictures of real objects were used. Our power estimation followed the following steps. First, we fitted a linear mixed-effects model on the RT data with both fixed (sentence type, picture type, and the interaction between them) and random effects (participants and items). Second, we estimated a power of 80 % on different sample sizes (from 50 to 120) with a *t*-value of 2 as a significance threshold. Third, we ran 1000 repetitions in the simulation process to test how many participants would be needed to detect a critical interaction of interest between sentences and pictures. Fourth, as per the suggestion of Kumle et al. (2021), we reduced all beta coefficients by 15 % to find the smallest effect size of interest. The simulation analyses suggested running an experiment with at least 80 participants to ensure $>$ 80 % power for the interaction.

Participants

In line with the power analysis, 81 university students ($M_{age} = 20.59$, $SD_{age} = 5.59$) participated in the experiment (of whom 70 were females) in exchange for course credit. In Experiments 1 to 2B, all participants were native speakers of European Portuguese.

Materials

Twenty-eight experimental sentence pairs were created describing either a minimal change of an object's state (e.g., "Jane *chose* a banana" [in original Portuguese language: "Jane escolheu uma banana"]) or a substantial change of an object's state (e.g., "Jane *stepped on* a banana" [in original Portuguese language: "Jane pisou uma banana"]). In addition, we created twenty-eight same-sized (385x385 pixels) experimental

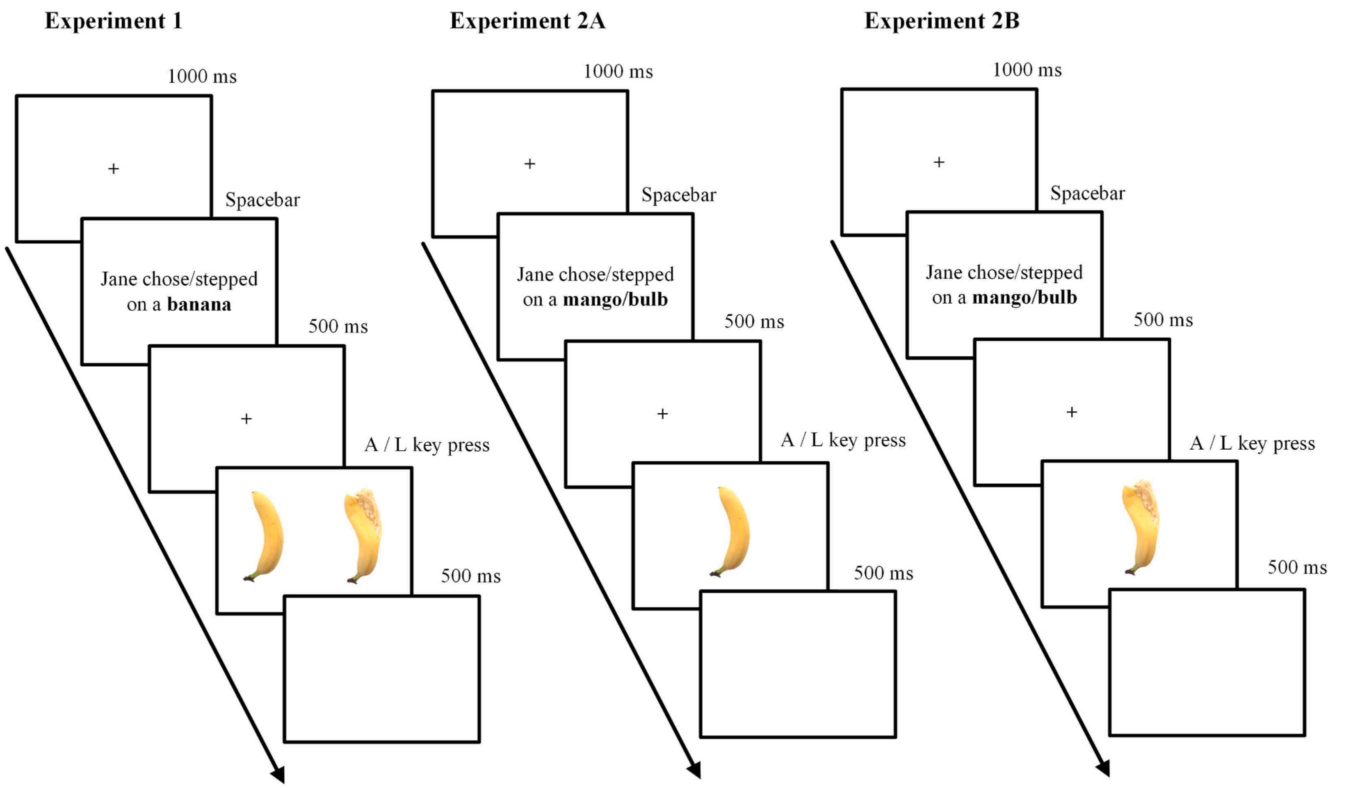

Fig. 2. Trial sequence in Experiments 1, 2A, and 2B Note. In all three experiments, each trial started with a black fixation cross displayed for 1000 ms (ms), followed by a sentence implying either a minimal or a substantial object change. When participants pressed the space bar to indicate that they finished reading, the sentence was replaced by a fixation cross (presented for 500 ms), followed by a pictured object that was either mentioned or not in the preceding sentence. Participants had to decide whether the object was mentioned in the sentence by pressing an L button for a "yes" response and an A button for a "no" response. After participants had indicated their decision, a blank screen replaced the picture (displayed for 500 ms), and a new trial began. The major difference between the experiments concerns the visual stimuli. Whereas in Experiment 1, target visual stimuli either matched or mismatched the state of the object implied by the sentence, in Experiments 2A and 2B, target visual stimuli were either semantically related or unrelated to the objects mentioned in the sentence. The only difference between Experiments 2A and 2B was in the state of the visual object displayed on the screen: intact and modified states in Experiments 2A and 2B, respectively.

image-pairs showing either the original state (e.g., intact banana) or the modified state (e.g., smashed banana) of the object (see Fig. 1, for sentence and picture samples). Finally, in addition to 28 experimental sentence-picture pairs requiring a "yes" response, 28 filler sentence-picture pairs were created requiring a "no" response (that is, the pictured objects mismatched the one from the sentence). Filler sentences had the same structure as experimental sentences. Half of the filler pictures depicted objects in the original state and the other half in a modified state. Finally, to motivate participants to read sentences attentively, half of the filler items were followed by comprehension questions, with participants providing an equal number of "yes" and "no" responses[1].

Design and procedure

Four lists of stimuli were created, and each experimental sentence-picture pair appeared in only one of the following conditions in each list: minimal sentence-original picture; minimal sentence-modified picture; substantial sentence-original picture; and substantial sentence-modified picture. Each participant was randomly assigned to each list and completed the task from one list only. This led to a 2 (sentence: minimal vs. substantial) × 2 (picture: original vs. modified) within-participants design.

The study was presented online via the web-based service Psytoolkit[2] (Stoet, 2010, 2017). The experiment began with six practice trials to allow participants to get used to the task. As shown in Fig. 2, each trial of the main part of the experiment started with a fixation cross in the middle of a screen (that lasted for 1000 ms). Then, a sentence appeared at the center of the screen until participants indicated that they had finished reading the sentence (by pressing the Spacebar button). Afterward, the sentence was replaced by a fixation cross (presented for 500 ms), followed by a pictured object (in either an original or modified state) that was either mentioned or not in the preceding sentence. Participants were asked to decide as fast as possible whether the pictured object had appeared in the preceding sentence by pressing an L button for a "yes" response and an A button for a "no" response[3]. After participants had indicated their decision, a blank screen replaced the picture (for 500 ms), and a new trial began. All trials were presented in random order.

Results

Data treatment and statistical analyses

Prior to analyses, and similar to previous research using a sentence-picture verification task (e.g., Connell, 2007; de Koning et al., 2017; Horchak & Garrido, 2021), in all three experiments, we only considered the responses of participants with accuracy rates equal to or higher than 80 %. This led to the removal of three, four, and six participants in Experiment 1, Experiment 2A, and Experiment 2B, respectively. For RT analyses, where the analyses were always performed on correct responses only, we removed responses faster than 300 ms and slower than 3000 ms, as well as responses with RTs 2.5 SDs higher than the relevant condition's mean. This trimming procedure eliminated 3.09 %, 3.99 %, and 4.66 % of the data in Experiment 1, Experiment 2A, and Experiment 2B, respectively. Finally, due to positive skewness, in all experiments, RTs were log-transformed to normalize the distribution.

All statistical analyses in Experiments 1 to 2B were conducted in the open-source programming language R (R Core Team, 2020). The dataset was analyzed using logistic (accuracy analyses) and linear (RT analyses) mixed-effects modeling. The full model[4] included sentence type (minimal change vs. substantial change), picture type (original state vs. modified state), and their interaction as fixed effects; and by-participant and by-item random intercepts. Additionally, the model also included by-participant random slopes for sentence type, picture type, as well as the interaction term. If the estimation procedure for the full random-effects specification resulted in non-convergence or overfitting warnings, we simplified the model's random-effects structure by first removing random correlations (see Barr et al., 2013, for further discussion). If this did not help, then we removed random slopes that had little influence on the results. The decision to simplify a "maximal" model was driven by the need to balance Type I error and power, given that "maximal" models lose power if their complexity is not supported by the data (Matuschek et al., 2017). The final random-effects structure of each model (for each experiment) used to report the results is specified in Tables 1, 3, and 4. The fixed effects predictors were sum-coded (-1, 1) to be able to obtain the main effects. If a significant interaction was observed, we performed follow-up analyses using the "emmeans" R package (Lenth, 2017) and selected the Holm method for controlling for family wise error rate (Aickin & Gensler, 1996).

Accuracy analyses

As shown in Table 1, the data showed no main effect of sentence type. However, picture type had a main effect, reflecting that the objects in modified state were verified less accurately ($M = 0.93$, $SD = 0.25$) than those in original state ($M = 0.97$, $SD = 0.18$). Furthermore, there was a significant interaction between sentences and pictures. Follow-up analyses demonstrated, as expected, that "modified" pictures were verified more accurately after the sentence implying a substantial change of state ($M = 0.96$, $SD = 0.19$) than a minimal ($M = 0.90$, $SD = 0.30$) change of state ($b = 1.058$, $SE = 0.277$, $z = 3.818$, $p < .001$). In contrast, "original" pictures were verified more accurately after the sentence implying a minimal change of state ($M = 0.99$, $SD = 0.12$) than a substantial ($M = 0.95$, $SD = 0.23$) change of state ($b = -1.458$, $SE = 0.415$, $z = -3.514$, $p < .001$).

RT analyses

The data of major interest are illustrated in Fig. 3. As can be seen in Table 1, the analyses showed no main effect of sentence type. However, as in the accuracy analyses, picture type had a main effect, with slower RTs for pictures in modified state ($M = 884$, $SD = 379$) than in the original state ($M = 803$, $SD = 326$). This result confirms previous findings that the original state has an advantage in response times relative to modified state (e.g., Horchak & Garrido, 2021; Kang et al., 2019). Furthermore, it is consistent with a well-established typicality effect whereby participants respond faster when seeing typical exemplars of a category (e.g., intact banana) than when seeing atypical (e.g., smashed banana) exemplars of a category (e.g., Rosch, 1975). Finally, there was also a significant interaction between sentences and pictures.

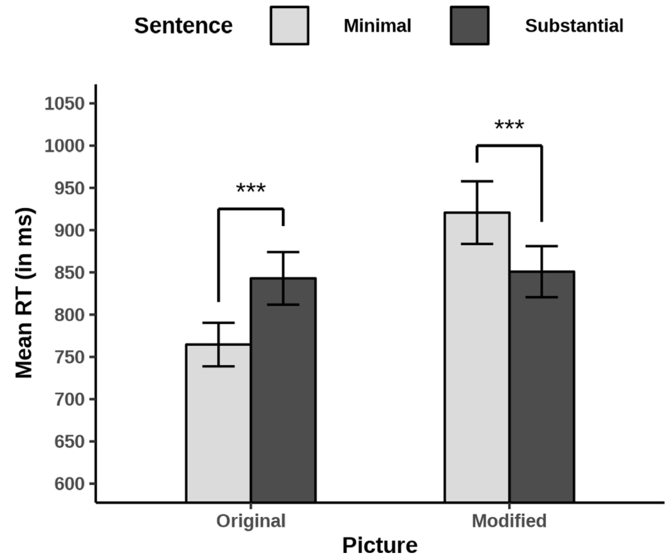

Fig. 3. Response Times in Experiment 1 Notes. ***p <.001.

Follow-up analyses revealed that pictures in modified state were verified faster after sentences implying a substantial state-change for an object than a minimal change ($b = -0.083$, $SE = 0.020$, $t = -4.136$, $p < .001$); and pictures in original state were verified faster after sentences implying a minimal state-change than a substantial one ($b = 0.087$, $SE = 0.020$, $t = 4.443$, $p < .001$). Furthermore, the segregation of the data by sentences revealed, as expected, that participants were faster to verify an "original" picture than a "modified" picture after reading a sentence suggesting a minimal change of state ($b = 0.179$, $SE = 0.023$, $t = 7.916$, $p < .001$). Crucially, however, the analyses showed that there was almost no difference in verification times between "original" ($M = 843$) and "modified" ($M = 851$) pictures after reading a sentence suggesting a substantial change of state ($b = 0.008$, $SE = 0.021$, $t = 0.395$, $p = .694$). Thus, this finding confirms the prediction that multiple distinct representations of an object are activated when representing object state-change in the substantial-change condition.

Experiment 2A

In Experiment 1, we established that appropriate object representations are modulated by the degree of implied state-change. Furthermore, we replicated previous findings (e.g., Hindy et al., 2012; Kang et al., 2019) that language comprehenders need to maintain various representations of an object in distinct states through which it has passed. However, as discussed before, language comprehenders may also be sensitive to multiple object state-representations of semantically related objects compared to semantically unrelated objects. That is, when a sentence implies a substantial change of object state (e.g., "Jane *stepped*

Table 1
Parameter Values for Fixed Effects in Mixed Logistic Regression Model (Analysis of Accuracy) and Mixed Linear Regression Model (Analysis of Log-transformed RTs) in Experiment 1.

Effect	Accuracy (GLMER)				RT (LMER)			
	b	SE	z	p	b	SE	t	p
Sentence	−0.100	0.125	−0.802	0.422	0.001	0.006	0.153	0.878
Picture	−0.506	0.205	−2.474	**0.013**	0.047	0.008	5.871	**<.001**
Sentence * Picture	0.629	0.125	5.041	**<.001**	−0.043	0.008	−5.659	**<.001**

Note. The final models used to report Accuracy and RT results included by-participant and by-item random intercepts. The major difference concerned random slopes. Due to overfitting problems, the final model to analyze Accuracy had to be simplified and only included a by-participant random slope for picture. The "maximal" model to analyze RTs converged successfully and had no overfitting problems. Numbers in bold refer to significant results.

on a mango"), participants may also have difficulty indicating that the subsequently presented object (e.g., a banana) was not mentioned in the sentence. We addressed this possibility in Experiment 2A.

To recap our predictions, if it is the case that after reading (1) "Jane stepped on a mango" multiple representations of the same object are activated, and after reading (2) "Jane chose a mango" only one object state is activated, then a semantically related banana, regardless of its pictured state, should get more priming from sentence (1) than from sentence (2), and hence elevated RTs. However, the amount of priming in sentence (1) should also vary continuously with the degree of change (as rated by different participants). Specifically, we expected to observe that the greater the difference between the state of the object implied by a "substantial state-change" sentence (1) and a "minimal state-change" sentence (2), the greater the conflict between two states would be. Consequently, the conflict between highly dissimilar object states (i.e., states that struggle to coexist) should have some inhibitory effect and, as a result, decrease the amount of priming afforded by sentence (1). This, in turn, should lead to the decreased RTs (i.e., less priming should make it easier for participants to reject the pictured banana).

As discussed before, no differences for semantically unrelated objects (e.g., bulb and banana) are expected because conceptual representations activated by the target word "bulb" only minimally overlap with those associated with the target object "banana". We tested these predictions in relation to objects depicted in their original (initial) states.

Method

Participants

Ninety-seven native-speaking university students (M_{age} = 20.43, SD_{age} = 4.48), of whom 82 were females, took part in the experiment in exchange for course credit.

Materials

As in Experiment 1, there were 28 experimental sentence-picture pairs and 28 filler sentence-picture pairs. However, there were some important differences. Whereas in Experiment 1, participants read sentences involving an object shown in the subsequent picture (thus requiring a "yes" response), in Experiment 2A, participants read sentences involving an object that was either semantically related (e.g., mango) or unrelated (e.g., bulb) to the subsequently presented (e.g., banana) object (thus requiring a "no" response). Half of the sentences described a semantically similar object, and the other half described a semantically unrelated object. Semantic similarity between objects from the sentences and objects from the pictures was calculated with word2vec word embeddings (e.g., Mikolov et al., 2013). Table 2 provides word2vec scores for targets used in each sentence-picture pair. Overall, the strength of semantic association was approximately three times higher between the concepts in the high similarity condition than in the low similarity condition.

Picture stimuli were the same as in Experiment 1, except that they all depicted an object in its original state (see Fig. 1, for sentence and picture samples). Finally, in addition to 28 experimental items requiring a "no" response, 28 filler sentence-picture pairs were created requiring a "yes" response (that is, the pictured objects matched the one from the sentence). Filler sentences had the same structure as experimental sentences. Half of the filler pictures depicted objects in the original state and the other half in a modified state.

Additionally, a separate sample of 72 native Portuguese participants (M_{age} = 35.05, SD_{age} = 9.49; 40 females), of whom 28 were university students, and 44 were freelancers on Clickworker, were asked to rate the degree of object change in each of the critical sentences (56 sentences). There were two lists of stimuli to ensure that each participant rated only one version of the event. For example, if participants from list 1 read the sentence "Jane chose a mango", then participants from list 2 read the

Table 2
Word2vec scores for word pairs in high- and low semantic similarity conditions.

#	Pictured target	Word competitor from sentence: Word Low similarity	Word competitor from sentence: Word High similarity	Word2vec scores: Word Low similarity	Word2vec scores: Word High similarity
1	banana	guitar	mango	0.099	0.637
2	egg	thermometer	cake	0.26	0.338
3	plate	TV	cup	−0.007	0.258
4	glasses	watermelon	bottle	0.185	0.515
5	iPhone	washbasin	iPad	0.143	0.757
6	light bulb	bowl	candle	0.191	0.518
7	teapot	laptop	mug	0.189	0.375
8	mirror	headphones	telescope	0.158	0.335
9	sushi	hot dog	shrimp	0.201	0.414
10	strawberry	sunglasses	raspberry	0.127	0.624
11	cracker	carrot	waffle	0.196	0.309
12	ice cream	vase	yogurt	0.255	0.363
13	tomato	cigarette	pepper	0.19	0.484
14	bottle	egg	jar	0.216	0.59
15	chocolate bar	toothpaste	cookie	0.253	0.433
16	wall clock	strawberry	painting	0.12	0.19
17	mug	projector	plate	0.158	0.219
18	cake	alarm clock	muffin	0.09	0.469
19	iPad	hamburger	iPhone	0.152	0.757
20	tile	radio	porcelain figure	0.03	0.362
21	windshield	pen	headlights	0.122	0.505
22	wine glass	fan	bowl	0.089	0.191
23	flan pudding	cardboard box	custard tart	0.253	0.715
24	mango	balloon	banana	0.095	0.637
25	donut	broccoli	bread	0.325	0.431
26	computer mouse	chair	keyboard	0.111	0.526
27	papaya	cigar	melon	0.137	0.513
28	rice cracker	calculator	flan pudding	−0.044	0.314
				0.153	**0.456**

Note. Numbers in bold indicate average values of semantic similarity in the relevant column. The word2vec scores are cosines (scale –1 to + 1) obtained via "Latent Semantic Analysis @ CU Boulder" website (http://wordvec.colorado.edu/).

Table 3
Parameter Values for Fixed Effects in Mixed Logistic Regression Model (Analysis of Accuracy) and Mixed Linear Regression Model (Analysis of Log-transformed RTs) in Experiment 2A.

Effect	Accuracy (GLMER)				RT (LMER)			
	b	SE	z	p	b	SE	t	p
Sentence	−0.066	0.146	−0.451	0.652	0.013	0.005	2.509	**0.012**
Similarity	−2.079	0.659	−3.156	**0.002**	0.076	0.005	14.913	**<.001**
Sentence * Similarity	−0.012	0.147	−0.079	0.937	0.011	0.005	2.247	**0.025**

Note. The final models used to report Accuracy and RT analyses included by-participant and by-item random intercepts. The major difference concerned random slopes. Due to overfitting problems, the final model to analyze Accuracy had to be simplified and only included a by-participant random slope for similarity. The final model to analyze RTs only included random intercepts (no slopes). Numbers in bold refer to significant results.

Table 4
Parameter Values for Fixed Effects in Mixed Logistic Regression Model (Analysis of Accuracy) and Mixed Linear Regression Model (Analysis of Log-transformed RTs) in Experiment 2B.

Effect	Accuracy (GLMER)				RT (LMER)			
	b	SE	z	p	b	SE	t	p
Sentence	−0.141	0.138	−1.016	0.310	0.021	0.006	3.794	**<.001**
Similarity	−1.718	0.476	−3.612	**<.001**	0.109	0.006	16.972	**<.001**
Sentence * Similarity	−0.047	0.138	−0.338	0.736	0.014	0.006	2.468	**0.014**

Note. The final models used to report Accuracy and RT analyses included by-participant and by-item random intercepts. The major difference concerned random slopes. Due to overfitting problems, the final models to analyze Accuracy and RTs had to be simplified and only included a by-participant random slope for similarity. Numbers in bold refer to significant results.

sentence “Jane stepped on a mango”. The instructions were identical to those used by Solomon et al. (2015): Participants were instructed to rate “the degree to which the depicted object will be at all different after the action occurs than it had been before the action occurred” using a scale from 1 (Just the same) to 7 (Completely changed). The mean change rating for items in the substantial state-change (e.g., the “step on” action) condition ($M = 5.68$, $SD = 0.74$) was significantly greater ($t\,(54) = -21.75$, $p < .001$) than the mean change rating for items in the minimal state-change (e.g., the “choose” action) condition ($M = 1.75$, $SD = 0.60$).

Design and procedure

Given the use of only one picture type (i.e., pictured object in original state), the experimental design was 2 (sentence: minimal change vs. substantial change) × 2 (semantic similarity: low vs. high). Like in Experiment 1, four lists of stimuli were created, and each experimental sentence-similarity pair appeared in the following conditions in each list: minimal sentence-low similarity; minimal sentence-high similarity; substantial sentence-low similarity; and substantial sentence-high similarity. Each participant was randomly assigned to each list and completed the task from only one list. The procedure was the same as in Experiment 1.

Results

Data treatment and statistical analyses

Data were treated in the same way as in Experiment 1. Statistical analyses were the same, except that we replaced a fixed effect of picture type with a fixed effect of semantic similarity in the final model (for both Accuracy and RT analyses).

Accuracy analyses

As shown in Table 3, the results showed that semantic similarity was the only significant main effect, reflecting the fact that participants were less accurate when semantic similarity between the objects from the sentence and the picture was high ($M = 0.91$, $SD = 0.28$) rather than low ($M = 0.99$, $SD = 0.11$). This effect fits the idea that participants experienced difficulties in resolving a competition between shared characteristics of the target and a semantically similar concept.

RT analyses

As demonstrated in Table 3, the results showed a main effect of sentence type, suggesting that participants were slower when they processed a sentence implying a substantial state-change ($M = 791$, $SD = 276$) rather than minimal state-change ($M = 773$, $SD = 265$). There was also a strong main effect of semantic similarity, reflecting that, as expected, high semantic similarity between sentence and picture items led to slower responses ($M = 846$, $SD = 298$) than low semantic similarity ($M = 724$, $SD = 228$). However, there was also an interaction between sentence type and the levels of semantic similarity.

As shown in Fig. 4 **(panel A)**, follow-up analyses showed that participants took more time to verify a picture in the substantial state-change ($M = 868$, $SD = 306$) condition (e.g., “smash” action) than in the minimal state-change ($M = 824$, $SD = 287$) condition (e.g., “observe” action) when semantic similarity between sentence and picture items was high ($b = 0.048$, $SE = 0.015$, $t = 3.279$, $p = .002$). However, there was no such difference when semantic similarity between sentence and picture items was low ($b = 0.003$, $SE = 0.014$, $t = 0.190$, $p = .849$). Thus, these results confirm our prediction that participants are sensitive to multiple state-change representations of semantically related objects.

To establish the effect of dissimilarity-based interference, we subtracted the mean change rating for the “substantial change” items (e.g., “Jane *stepped* on a mango”) from the mean change rating for the “minimal change” items (e.g., “Jane *chose* a banana”) and then used this difference score for the analysis[5]. The Pearson correlation[6] between the rated degree of object change (difference score) and participants’ RTs in the “substantial change” condition was significant ($r\,(560) = -0.096$, $p = .022$). As can be seen from Fig. 4 (panel B), greater dissimilarity between object states led to the reduction of participants’ RTs. This is consistent with the prediction that highly dissimilar object states mutually inhibit one another.

Experiment 2B

The results of Experiment 2A showed that the verification of pictures depicting objects in their original state was longer after participants read

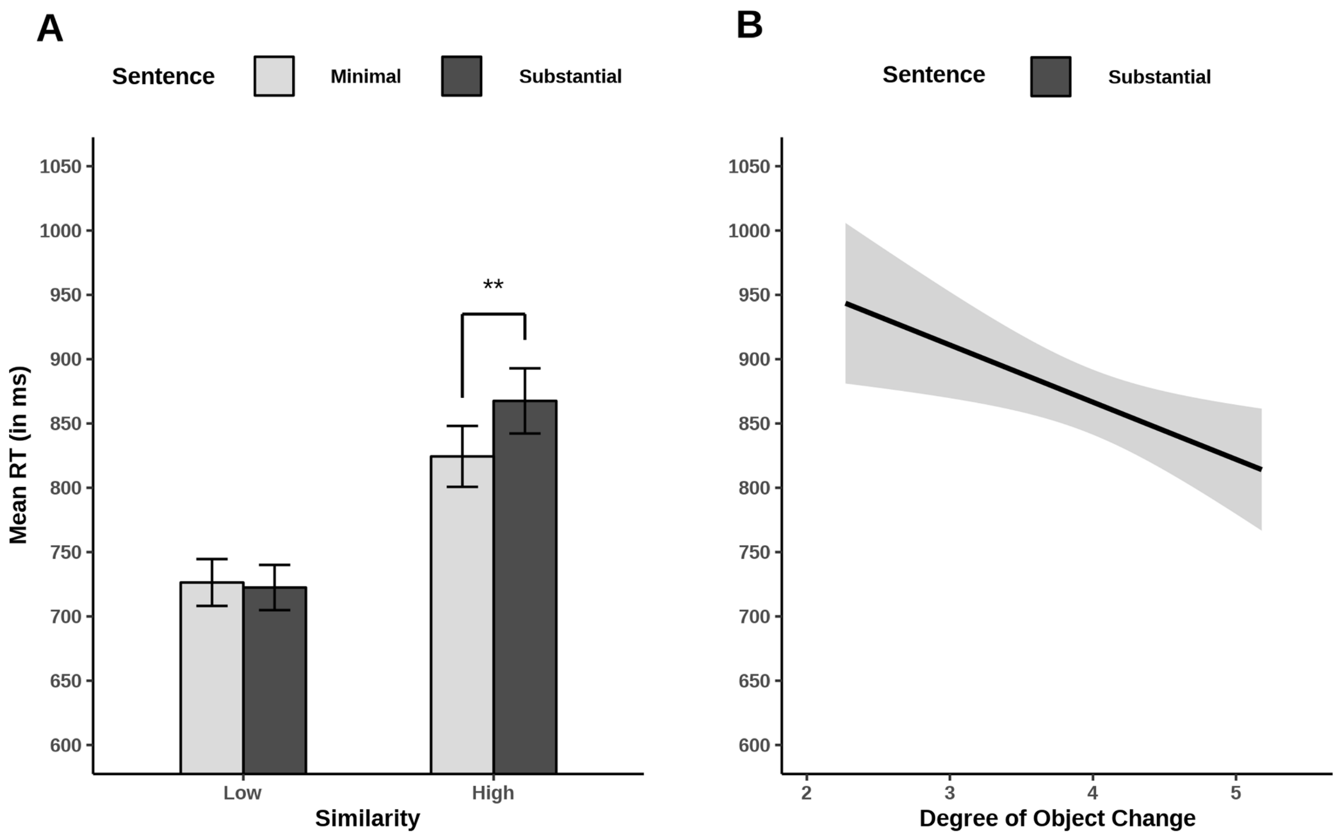

Fig. 4. Response Times in Experiment 2A With Pictured Objects in Original State Note. (A) Participants' responses in each of the critical conditions. Error bars reflect 95 % confidence intervals. **p <.01. (B) Item analysis for high similarity condition using the degree of object change (difference score) as a continuous predictor. Specifically, the line of best fit reflects how participants' RTs in the "substantial" sentence condition change as a function of the difference between the state of the object implied by a "substantial state-change" sentence (e.g., "Jane stepped on a mango") and a "minimal state-change" sentence (e.g., "Jane chose a mango"). The gray shading around the line represents the 95 % confidence interval around the line of best fit.

a sentence like (1) "Jane *stepped* on a mango" than (2) "Jane *chose* a mango". Our interpretation of this finding is that sentence (1) requires the cognitive system to maintain multiple incompatible representations of an object (original and modified), which, in turn, leads to slower verification times. However, it could be argued that after reading sentence (1), participants maintained only a single representation of an object in its final modified state (that is, the original state was successfully overwritten during comprehension). Consequently, response times were different just because participants found it easier to match a pictured object in its prototypical state (banana) with the intact representation of an object (mango) in sentence (2). Indeed, prior research suggests that object typicality may affect the saliency of the object-state representation (e.g., Nosofsky, 1988). At the same time, if there were indeed more priming for the depicted "banana" after the more typical representation of a mango in "Jane *chose* a mango", then it would be more reasonable to predict the opposite of what we observed in Experiment 2A. Specifically, we should have observed longer "no" responses after participants read "Jane *chose* a mango" (which should lead to a more typical representation of a banana) than after "Jane *stepped* on a mango" (which should lead to a more atypical representation of a banana).

Whatever the correct explanation, it is important to entirely rule out the concern that the state of the depicted object may have any significant effect on the observed results. To this end, in Experiment 2B, we used the same stimuli and procedure as in Experiment 2A, except that participants saw a pictured object in a modified state (e.g., squashed banana). If distinct representations of an object are activated and maintained (e.g., intact mango vs. squashed mango) during sentence reading, we should observe the same results as in Experiment 2A: longer picture verification latencies after sentences implying a substantial change in an object's state.

Method

Participants

Ninety-five native-speaking university students ($M_{age} = 20.51$, $SD_{age} = 4.65$), of whom 78 were females, took part in the experiment in exchange for course credit.

Materials

Materials were the same as in Experiment 2A, except that experimental picture stimuli depicted an object in its modified state (see Fig. 1, for a sample).

Design and procedure

The design and procedure were the same as in Experiment 2A.

Results

Data treatment and statistical analyses

Data treatment and statistical analyses were the same as in Experiment 2A.

Accuracy analyses

Similar to Experiment 2A, and as shown in Table 4, the results showed that semantic similarity was the only significant main effect, reflecting the fact that participants were less accurate when there was a high ($M = 0.91$, $SD = 0.29$) rather than low ($M = 0.99$, $SD = 0.12$) semantic similarity between the objects from the sentence and the picture.

RT analyses

The results showed a main effect of sentence type, suggesting that participants were slower when they processed a sentence implying a substantial change of state ($M = 852$, $SD = 323$) rather than a minimal change of state ($M = 823$, $SD = 306$). There was also a strong main effect of semantic similarity, reflecting the fact that, as in Experiment 2A, high semantic similarity between sentence- and picture items led to longer verification latencies ($M = 940$, $SD = 357$) than low semantic similarity ($M = 746$, $SD = 236$). Finally, there was also an interaction between sentence type and the levels of semantic similarity.

In line with our prediction, and as shown in Fig. 5 **(panel A)**, follow-up analyses revealed that participants were slower to verify a "modified" picture in the substantial change-state ($M = 968$, $SD = 366$) condition (e.g., "step" action) than the minimal change-state ($M = 912$, $SD = 347$) condition (e.g., "choose" action) when semantic similarity between sentence and picture items was high ($b = 0.070$, $SE = 0.016$, $t = 4.304$, $p < .001$). However, there was no such difference when semantic similarity between sentence and picture items was low ($b = 0.015$, $SE = 0.015$, $t = 0.965$, $p = .335$). Thus, these results once again confirm our hypothesis that multiple representations of an object compete when representing object state-change in a sentence like "Jane *stepped on* a mango"; and that comprehenders are sensitive to states of semantically related objects.

Finally, the Pearson correlation between the rated degree of object change (difference score) and participants' RTs in the "substantial change" condition was significant ($r\,(518) = -0.087$, $p = .048$). As can be seen from Fig. 5 (panel B), greater dissimilarity between object states led to the reduction of participants' RTs. This replicates the result we observed in Experiment 2A and lends further credence to our argument that highly dissimilar object states mutually inhibit one another.

Ruling out visual similarity as an explanation for observed effects

The focus of the current research was on the effect of state-change on semantically similar items. However, one may wonder if visual similarity between the items could explain some of our observed effects. This issue is important to address due to the compelling body of evidence supporting the impact of visual similarity on language comprehension (e.g., Cooper, 1974; Huettig et al., 2004), juxtaposed with the equally noteworthy top-down influences of semantic relatedness on visual processing (e.g., Cree & McRae, 2003; de Groot et al., 2016).

Considering the evidence above, we obtained the norming data for the visual and semantic similarity of our stimuli. Our specific goals were the following. First, we wanted to replicate the result of de Groot (2006) that participants cannot fully separate semantically similar and visually similar items compared to unrelated ones. Second, given the focus of the paper on semantic similarity, we wanted to confirm that our items from the high-similarity condition are rated higher on semantic similarity than visual similarity. Third, we wanted to identify the object pairs that were high on visual similarity (i.e., above the scale mid-point) to test whether the results would replicate without these items. In other words,

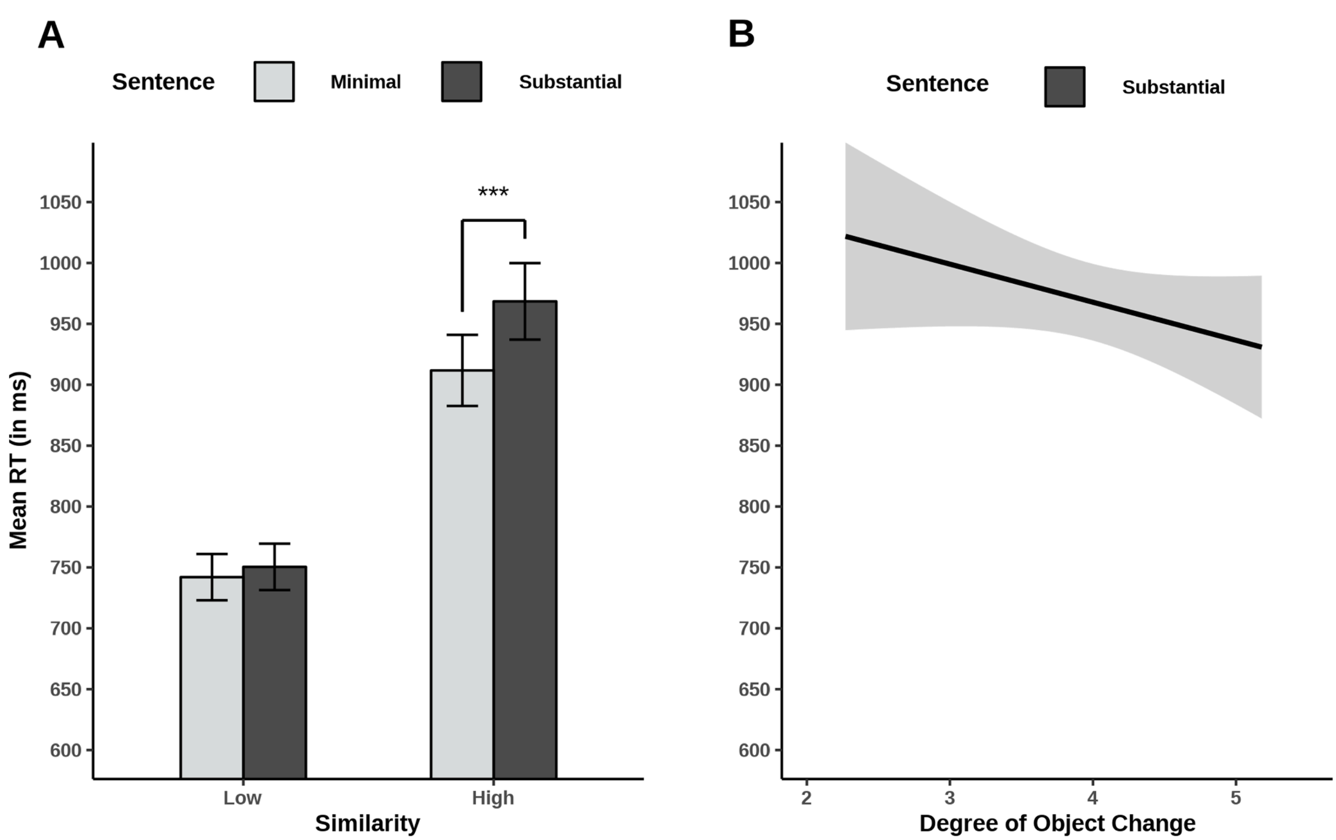

Fig. 5. Response Times in Experiment 2B With Pictured Objects in Modified State Note. (A) Participants' responses in each of the critical conditions. Error bars reflect 95 % confidence intervals. ***p <.001. (B) Item analysis for high similarity condition using the degree of object change (difference score) as a continuous predictor. Specifically, the line of best fit reflects how participants' RTs in the "substantial" sentence condition change as a function of the difference between the state of the object implied by a "substantial state-change" sentence (e.g., "Jane stepped on a mango") and a "minimal state-change" sentence (e.g., "Jane chose a mango"). The gray shading around the line represents the 95 % confidence interval around the line of best fit.

we wanted to ensure our results do not depend on these particular items rated high on visual similarity.

To achieve the above goals, 72 native-speaking Portuguese participants[7] saw word pairs (one word pair at a time) belonging to either the low similarity condition or the high-similarity condition and were asked to rate them on visual and semantic relatedness using a 7-point scale (1 = Not similar at all; 7 = Highly similar). Two lists of stimuli were created, and each participant was randomly assigned to each list. This

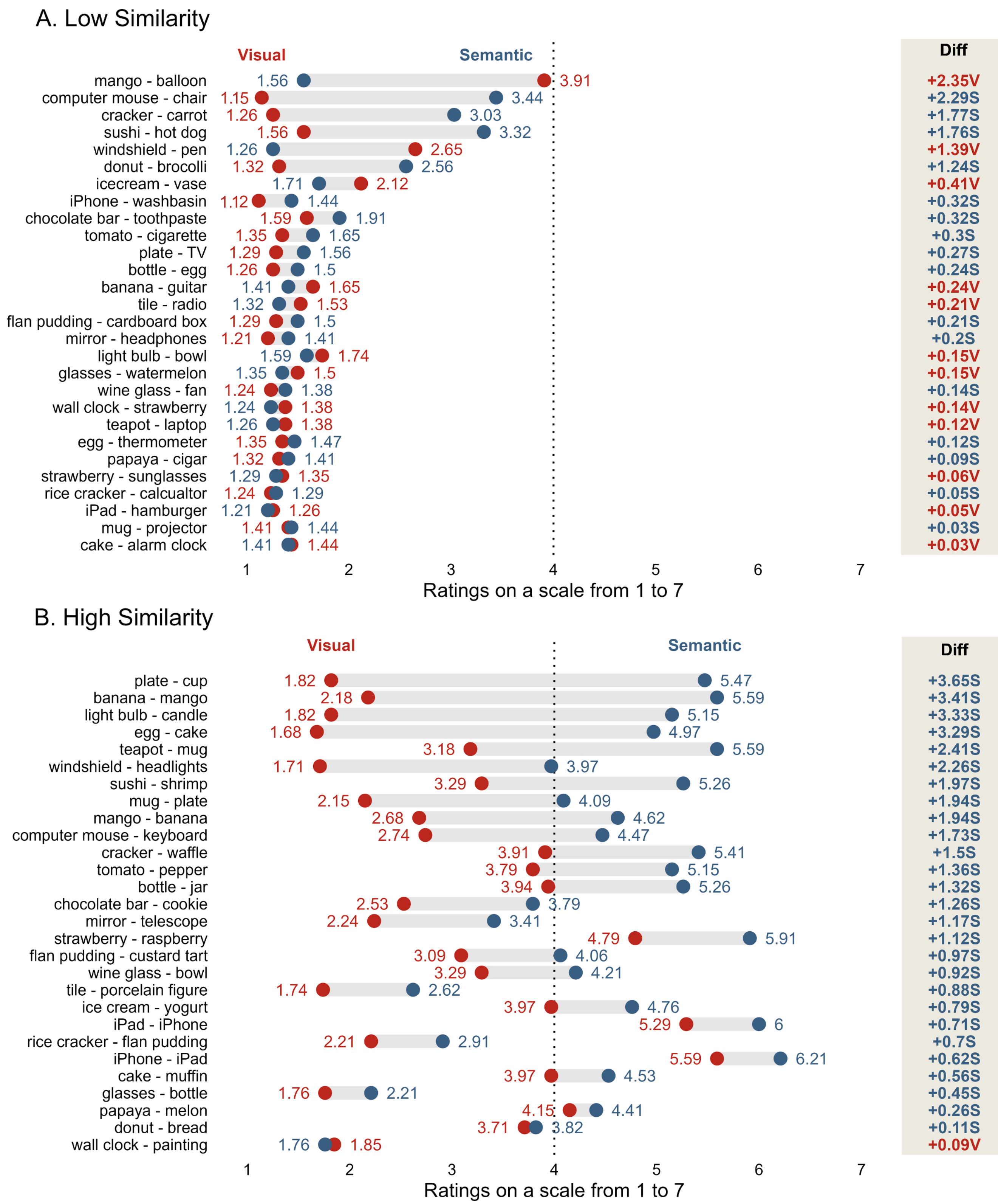

Fig. 6. Ratings of Visual and Semantic Similarity in High and Low Similarity ConditionsNote. The dashed line represents a scale mid-point. The column called "Diff" represents the difference in ratings for visual (V) and semantic (S) similarities. The items are ordered in ascending order (from bottom to top) based on the difference in ratings for visual similarity and semantic similarity.

was done to ensure that each survey participant rated only one version of each similarity type (i.e., high or low). For instance, participants from list 1 rated the similarity between a mango and a banana, while participants from list 2 rated the similarity between a guitar and a banana. We used the same instructions as in a study of Jiang et al. (2022) to define these two types of similarities. For visual similarity, we asked participants to rate how visually similar the two objects are based on how they **look alike** (e.g., similar shape, size, color, etc.). For participants to understand better what we mean by visual similarity, we asked them to consider the similarity between balloons and watermelons as an example. Specifically, we explained that balloons and watermelons may be rarely used together, but they are visually similar given their rounded shape. For semantic similarity, we asked participants how semantically similar the objects are based on how much they **have to do with each other**. Participants were told that two objects could be judged to be semantically similar if they (a) are often used together or complement each other (e.g., hammer and nail); serve the same purpose (e.g., fork and spoon); often occur in the same situation or environment (e.g., desk and computer); or are exemplars of the same category (e.g., guitar and piano).

As demonstrated in Fig. 6, the data showed that there was substantial agreement between corpus-based measures (see Table 1) and human judgments as to which object pairs have low and high similarity. Similar to de Groot et al. (2006), we found that ratings of visual similarity were significantly higher in the high similarity (M = 3.04, SD = 1.14) condition than in the low similarity (M = 1.53, SD = 0.56) condition, t (54) = -6.283, p <.001. Likewise, ratings of semantic similarity were significantly higher in the high similarity (M = 4.49, SD = 1.14) condition than in the low similarity (M = 1.68, SD = 0.62) condition, t (54) = -11.43, p <.001. However, ratings of semantic similarity were significantly higher than ratings of visual similarity in the high similarity condition, t (54) = − 4.748, p <.001. This suggests that semantic and visual relationships are clearly distinguishable within our set.

At the same time, participants' ratings of visual similarity revealed four object pairs (strawberry-raspberry, iPad-iPhone, iPhone-iPad, papaya-melon) that were rated high on visual similarity (i.e., above the scale mid-point). To rule out the concern that items with high visual similarity could explain the observed pattern of results, we reanalyzed the data from Experiments 2A and 2B by excluding these items. As can be seen from Table 5, the critical interaction of interest between sentence type (minimal vs. substantial) and similarity type (low vs. high) was observed after excluding visually similar items. With regard to Experiment 2A (with intact pictured objects), follow-up analyses showed that responses in the substantial change-state condition were slower than in the minimal change-state when semantic similarity between sentence and picture items was high (b = 0.048, SE = 0.016, t = 3.046, p =.005). However, there was no such difference when semantic similarity between sentence and picture items was low (b = -0.002, SE = 0.015, t = -0.139, p =.889). The same pattern was observed in Experiment 2B (with modified pictured objects). Responses were slower in the substantial change-state condition than in the minimal change-state condition when semantic similarity between the items was high (b = 0.068, SE = 0.017, t = 3.969, p <.001). However, there was no difference for object pairs with low similarity (b = 0.016, SE = 0.016, t = 0.963, p =.335). Thus, these results replicate all the critical effects observed before and suggest that high visual similarity had no substantial effect on the findings.

General discussion

In this research, we reported the findings of three experiments that explored whether multiple representations of objects in their different states are activated during language comprehension. Experiment 1 conceptually replicated the results of Kang et al. (2019) with a different stimulus set. More specifically, we found that objects in the modified state were verified faster when a sentence implied a substantial object state-change rather than a minimal state-change. In contrast, objects in the original state were verified faster when a sentence implied a minimal (or no change) object state-change than a substantial state-change. Finally, when segregating the data by sentences, we also found that verification latencies of pictures depicting an original state of an object (i.e., intact banana) were faster than verification latencies of pictures depicting a modified state of an object (i.e., squashed banana) when the sentence described a minimal change of state ("He *chose* a banana"). Importantly, however, there were no differences in verification latencies when the sentence described a substantial change of state ("He *stepped* on a banana"). This confirms the hypothesis that language comprehenders maintain multiple representations of an object in various states in the substantial-change condition.

Experiments 2A and 2B extended the findings of Experiment 1 in several important ways. First, when participants had to indicate that a pictured object was not mentioned in the preceding sentence (e.g., banana), their verification latencies were slowed down by semantically related items mentioned in the sentence (e.g., mango) rather than semantically unrelated items (e.g., bulb), but only after a sentence like (1) "Jane *stepped* on a mango" than a sentence like (2) "Jane *chose* a mango". This suggests that comprehenders found it challenging to inhibit multiple representations of the same object (e.g., original vs. modified mango) in sentence (1).

Finally, the present research is also consistent with previous findings (e.g., Hindy et al., 2012; Solomon et al., 2015) suggesting that state ambiguity engenders representational conflict. Specifically, and in line with this prior research, our data show that mutual inhibition of object states is not so much due to the action word itself but rather to the extent to which the object changes in state as implied by action and the physical properties of the object. As discussed before, stepping on a papaya is not the same as stepping on a mango, for example, as the latter offers more resistance and might not flatten as easily. This means that the alternative states of a mango would be more similar after reading "Jane stepped on vs. chose a mango" than the alternative states of a papaya after reading "Jane stepped on vs. chose a papaya". This gives us insight into exactly when we could expect interference between multiple instantiations of the same object. We conjecture that the more similar the alternative object states (e.g., different states of the mango), the greater the amount of semantic priming that a semantically related object (e.g., banana) receives. To some extent, this proposal is consistent with the findings of Connell and Lynott (2009) about object color. Specifically, they suggested that context-implied information (e.g., white bear) is held in parallel with the more typical information about an object (e.g., brown bear). However, our results suggest that the

Table 5
Parameter Values for Fixed Effects in Mixed Linear Regression Model (Analysis of log-transformed RTs) When Excluding Visually Similar Items in Experiments 2A and 2B.

Effect	Experiment 2A (intact object)				Experiment 2B (modified object)			
	b	SE	t	p	b	SE	t	p
Sentence	0.011	0.005	2.087	**0.040**	0.021	0.006	3.546	**<.001**
Similarity	0.064	0.005	11.969	**<.001**	0.093	0.007	13.289	**<.001**
Sentence * Similarity	0.012	0.005	2.330	**0.020**	0.013	0.006	2.218	**.027**

Note. Numbers in bold refer to significant results.

effects of state-change on event comprehension could be more nuanced than the study of Connell and Lynott (2009) suggests. Specifically, our data demonstrate that the more dissimilar the alternative object states, the less likely it is that the atypical and typical representations could be successfully held in parallel. On the contrary, after reading a sentence like "Jane stepped on a papaya", the highly dissimilar object states may mutually inhibit one another, and, as a result, the amount of semantic priming that a semantically related object (e.g., banana) receives should significantly decrease (i.e., participants need less time to reject the pictured banana). Future work is required to say with any precision exactly how much it can decrease.

These findings break new ground in studies of event comprehension through the prism of the IOH account (Altmann & Ekves, 2019). Theoretically, they show that language comprehenders are sensitive to state changes of semantically related objects (compared to previous research showing the sensitivity of language comprehenders to state changes of target objects). On a methodological level, the present research is the first one that addressed the question of how/why object state-change representation should occur when one needs to compare the wrong object (i.e., provide a "No" response) to the one mentioned in a sentence in a sentence-picture verification task (as opposed to previous research focusing on what happens when one needs to positively verify the object to the sentence). Finally, the reported data reframe and contextualize a robust finding from prior research studying the organization of semantic knowledge, namely that the interference effects arise due to semantic similarity between words and pictured targets (e.g., Schriefers et al., 1990; Vigliocco et al., 2004). While in agreement with these studies, the present research goes considerably beyond this evidence as it sheds light on the complexity of interaction between both *similarity-based* interference (i.e., semantic overlap between items) and *dissimilarity-based* interference (i.e., competition between alternative object states). More precisely, our results suggest that the representation of an object arises through associations between the object and the others with which it tends to co-occur and through the representation of the trajectory of states of different (semantically) related objects coming together in space and time, whereby changes in the state of one are bound to changes in the state of another. Thus, interference effects may occur due to an increased number of semantic features in conflict between multiple state-changes of any given semantically related object.

By widening the lens towards the nature of conceptual processing, our results support the assumption of the IOH account (Altmann & Ekves, 2019) that comprehending events involves constructing dynamic representations of intersecting object histories (i.e., the initial and final states of the objects). Thus, our results are compatible with theories and models emphasizing dynamic views of cognition and conceptual flexibility (e.g., Kiefer & Pulvermüller, 2012; Tabor et al., 1997). Most broadly, the reported data cohere with empirical findings on the dynamics of interaction between episodic knowledge and semantic memory, which show that linguistic input provides a strong constraint on the activation of conceptual properties (e.g., Altmann & Kamide, 2007; Knoeferle & Crocker, 2007; Mirković & Altmann, 2019).

In conclusion, the experiments reported in this paper provide initial evidence that participants are sensitive to multiple state-changes of semantically similar objects and add to a growing body of evidence on the importance of object state-change representation during language comprehension (e.g., Kang et al., 2020; Misersky et al., 2019).

Footnotes

1. Comprehension questions were of no theoretical interest to us. We used them to motivate participants to pay attention to all the information provided in the sentence (rather than just the keywords). All participants had an accuracy of 50 % in answering these questions.

2. Previous research showed that the response measurement timing of PsyToolkit is reliable and comparable in its accuracy to the stimulus presentation software E-Prime 3 in a complex psycholinguistic task (Kim et al., 2019).

3. It is important to stress that participants were not questioned about the consequence of the action implied by the sentence. Both the instructions in the training phase and the experimental phase informed participants that they had to merely decide whether or not the pictured object was mentioned in the sentence.

4. Accuracy data were analyzed with generalized linear model (family binomial) using the following syntax: *Accuracy* ∼ *sentence* * *picture* + *(1* + *sentence* * *picture* | *ppt)* + *(1* | *item)*. RT data were analyzed with linear mixed model (REML procedure) using the following syntax: *RT* ∼ *sentence* * *picture* + *(1* + *sentence* * *picture* | *ppt)* + *(1* | *item)*.

5. We did not use the actual (not difference score) mean change ratings that participants provided in the substantial-change condition for one simple reason: across the different examples of both minimal and substantial change conditions the actual degree to which the object underwent change varied.

6. As all analyses of RTs in the present paper were performed on unaggregated data, we also retained participant-level data in the correlation analyses for the purpose of consistency. The analysis by items (i.e., when collapsing across participants for each item) similarly showed that greater dissimilarity between object states led to the reduction of RTs. However, the correlation between the rated degree of change and RTs was significant in Experiment 2A (r (26) = − 0.388, p =.041) but not in Experiment 2B (r (26) = − 0.225, p =.250).

7. Participants were the same as those used in the state-change rating task.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The materials, data, and R scripts for all experiments have been made available on OSF: https://osf.io/hxv2e/?view_only=db807f235f83415fbc1d7f210262e372

Acknowledgments

This research was supported by the Portuguese Foundation for Science and Technology (FCT) with a grant 2021.01551.CEECIND awarded to the first author. There are no conflicts of interest to disclose.

References

Aickin, M., & Gensler, H. (1996). Adjusting for multiple testing when reporting research results: The Bonferroni vs Holm methods. *American Journal of Public Health, 86*(5), 726–728. https://doi.org/10.2105/ajph.86.5.726

Altmann, G. T. M., & Ekves, Z. (2019). Events as intersecting object histories: A new theory of event representation. *Psychological Review, 126*, 817–840. https://doi.org/10.1037/rev0000154

Altmann, G. T. M., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language, 57*(4), 502–518. https://doi.org/10.1016/j.jml.2006.12.004

Banich, M., Milham, M., Atchley, R., Cohen, N., Webb, A., Wszalek, T., et al. (2000). fMRI studies of Stroop tasks reveal unique roles of anterior and posterior brain systems in attentional selection. *Journal of Cognitive Neuroscience, 12*, 988–1000. https://doi.org/10.1162/08989290051137521

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 10.1016/j.jml.2012.11.001. 10.1016/j.jml.2012.11.001.

Connell, L. (2007). Representing object colour in language comprehension. *Cognition, 102*(3), 476–485. https://doi.org/10.1016/j.cognition.2006.02.009

Connell, L., & Lynott, D. (2009). Is a bear white in the woods? Parallel representation of implied object color during language comprehension. *Psychonomic Bulletin & Review, 16*, 573–577. https://doi.org/10.3758/PBR.16.3.573

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language. A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology, 6*, 813–839. https://doi.org/10.1016/0010-0285(74)90005-X

Cree, G. S., & McRae, K. (2003). Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello (and many other such concrete nouns). *Journal of Experimental Psychology: General, 132*(2), 163–201. https://doi.org/10.1037/0096-3445.132.2.163
Cree, G. S., McRae, K., & McNorgan, C. (1999). An attractor model of lexical conceptual processing: Simulating semantic priming. *Cognitive Science, 23*(3), 371–414. https://doi.org/10.1016/S0364-0213(99)00005-1
Damian, M. F., & Bowers, J. S. (2003). Locus of semantic interference in picture-word interference tasks. *Psychonomic Bulletin & Review, 10*(1), 111–117. https://doi.org/10.3758/BF03196474
de Groot, F., Koelewijn, T., Huettig, F., & Olivers, C. N. L. (2016). A stimulus set of words and pictures matched for visual and semantic similarity. *Journal of Cognitive Psychology, 28*(1), 1–15. https://doi.org/10.1080/20445911.2015.1101119
Estes, Z., Golonka, S., & Jones, L. L. (2011). Thematic thinking: The apprehension and consequences of thematic relations. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 249–294). Elsevier Academic Press.
de Koning, B. B., Wassenburg, S. I., Bos, L. T., & van der Schoot, M. (2017). Mental simulation of four visual object properties: Similarities and differences as assessed by the sentence–picture verification task. *Journal of Cognitive Psychology, 29*, 420–432. https://doi.org/10.1080/20445911.2017.1281283
Gernsbacher, M. A., Goldsmith, H. H., & Robertson, R. R. W. (1992). Do Readers Mentally Represent Characters' Emotional States? *Cognition & Emotion, 6*(2), 89–111. https://doi.org/10.1080/02699939208411061
Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton MiZin.
Glenberg, A. M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language, 26*, 69–83. https://doi.org/10.1016/0749-596X(87)90063-5
Golonka, S., & Estes, Z. (2009). Thematic relations affect similarity via commonalities. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 1454–1464. https://doi.org/10.1037/a0017397
Hindy, N. C., Altmann, G. T. M., Kalenik, E., & Thompson-Schill, S. L. (2012). The effect of object state-changes on event processing: Do objects compete with themselves? *The Journal of Neuroscience, 32*(17), 5795–5803. https://doi.org/10.1523/JNEUROSCI.6294-11.2012
Horchak, O. V., & Garrido, M. V. (2021). Dropping bowling balls on tomatoes: Representations of object state-changes during sentence processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 47*(5), 838–857. https://doi.org/10.1037/xlm0000980
Horchak, O. V., & Garrido, M. V. (2022). Simulating background settings during spoken and written sentence comprehension. *Psychonomic Bulletin & Review, 29*(4), 1426–1439. https://doi.org/10.3758/s13423-022-02061-9
Horton, W. S., & Rapp, D. N. (2003). Out of sight, out of mind: Occlusion and the accessibility of information in narrative comprehension. *Psychonomic Bulletin & Review, 10*(1), 104–110. https://doi.org/10.3758/BF03196473
Huettig, F., & Altmann, G. T. M. (2004). The online processing of ambiguous and unambiguous words in context: Evidence from head-mounted eye-tracking. In M. Carreiras, & C. Clifton (Eds.), *The on-line study of sentence comprehension: Eyetracking, ERP and beyond* (pp. 187–207). New York, NY: Psychology Press.
Huettig, F., & Altmann, G. T. M. (2005). Word meaning and the control of eye fixation: Semantic competitor effects and the visual world paradigm. *Cognition, 96*(1), B23–B32. https://doi.org/10.1016/j.cognition.2004.10.003
Huettig, F., Quinlan, P. T., McDonald, S. A., & Altmann, G. T. M. (2006). Models of high-dimensional semantic space predict language-mediated eye movements in the visual world. *Acta Psychologica, 121*(1), 65–80. https://doi.org/10.1016/j.actpsy.2005.06.002
Jiang, Z., Sanders, D. M. W., & Cowell, R. A. (2022). Visual and semantic similarity norms for a photographic image stimulus set containing recognizable objects, animals and scenes. *Behavior Research Methods, 54*(5), 2364–2380. https://doi.org/10.3758/s13428-021-01732-0
Kang, X., Eerland, A., Joergensen, G. H., Zwaan, R. A., & Altmann, G. T. M. (2020). The influence of state change on object representations in language comprehension. *Memory & Cognition, 48*(3), 390–399. https://doi.org/10.3758/s13421-019-00977-7
Kiefer, M., & Pulvermüller, F. (2012). Conceptual representations in mind and brain: Theoretical developments, current evidence and future directions. *Cortex, 48*, 805–825. 10.1016/j.cortex.2011.04.006.
Kim, J., Gabriel, U., & Gygax, P. (2019). Testing the effectiveness of the Internet-based instrument PsyToolkit: A comparison between web-based (PsyToolkit) and lab-based (E-Prime 3.0) measurements of response choice and response time in a complex psycholinguistic task. *PLOS ONE, 14*(9), e0221802.
Knoeferle, P., & Crocker, M. W. (2007). The influence of recent scene events on spoken comprehension: Evidence from eye movements. *Journal of Memory and Language, 57*(4), 519–543. https://doi.org/10.1016/j.jml.2007.01.003
Kumle, L., Võ, M. L.-H., & Draschkow, D. (2018). Mixedpower: a library for estimating simulation-based power for mixed models in R. 10.5281/zenodo.1341047.
Kumle, L., Võ, M.-L.-H., & Draschkow, D. (2021). Estimating power in (generalized) linear mixed models: An open introduction and tutorial in R. *Behavior Research Methods, 53*(6), 2528–2543. https://doi.org/10.3758/s13428-021-01546-0
Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Harvard University Press.
Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 104*, 211–240. https://doi.org/10.1037/0033-295X.104.2.211
Lee, S. H., & Kaiser, E. (2021). Does hitting the window break it?: Investigating effects of discourse-level and verb-level information in guiding object state representations. *Language, Cognition and Neuroscience, 36*(8), 921–940. https://doi.org/10.1080/23273798.2021.1896013
Lenth, R. (2017). emmeans: Estimated Marginal Means, aka Least-Squares Means. *R package version*, (9), 1. https://CRAN.R-project.org/package=emmeans.
Lin, E. L., & Murphy, G. L. (2001). Thematic relations in adults' concepts. *Journal of Experimental Psychology. General, 130*, 3–28. https://doi.org/10.1037/0096-3445.130.1.3
MacLeod, C. M., & MacDonald, P. A. (2000). Interdimensional interference in the Stroop effect: Uncovering the cognitive and neural anatomy of attention. *Trends in Cognitive Sciences, 4*(10), 383–391. https://doi.org/10.1016/S1364-6613(00)01530-8
Mahon, B. Z., Costa, A., Peterson, R., Vargas, K. A. A., & Caramazza, A. (2007). Lexical selection is not by competition: A reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(3), 503–535. https://doi.org/10.1037/0278-7393.33.3.503
Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language, 94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001
McRae, K., & Boisvert, S. (1998). Automatic semantic similarity priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(3), 558–572. https://doi.org/10.1037/0278-7393.24.3.558
Mervis, C. B., & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology, 32*, 89–115. https://doi.org/10.1146/annurev.ps.32.020181.000513
Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems* (pp. 3111–3119).
Mirković, J., & Altmann, G. T. M. (2019). Unfolding meaning in context: The dynamics of conceptual similarity. *Cognition, 183*, 19–43. https://doi.org/10.1016/j.cognition.2018.10.018
Mirman, D., & Graziano, K. M. (2012). Individual differences in the strength of taxonomic versus thematic relations. *Journal of Experimental Psychology: General, 141*(4), 601. https://doi.org/10.1037/a0026451
Mirman, D., Landrigan, J.-F., & Britt, A. E. (2017). Taxonomic and thematic semantic systems. *Psychological Bulletin, 143*(5), 499–520. https://doi.org/10.1037/bul0000092
Mirman, D., & Magnuson, J. S. (2009). Dynamics of activation of semantically similar concepts during spoken word recognition. *Memory & Cognition, 37*(7), 1026–1039. https://doi.org/10.3758/MC.37.7.1026
Misersky, J., Slivac, K., Hagoort, P., & Flecken, M. (2021). The state of the onion: Grammatical aspect modulates object representation during event comprehension. *Cognition, 214*, Article 104744. https://doi.org/10.1016/j.cognition.2021.104744
Nozari, N., & Pinet, S. (2020). A critical review of the behavioral, neuroimaging, and electrophysiological studies of co-activation of representations during word production. *Journal of Neurolinguistics, 53*, Article 100875. https://doi.org/10.1016/j.jneuroling.2019.100875
Nosofsky, R. M. (1988). Exemplar-based accounts of relations between classification, recognition, and typicality. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(4), 700. https://doi.org/10.1037/0278-7393.14.4.700
R Core Team. (2020). R: A language and environment for statistical computing. R Foundation for Statistical. *Computing.* https://www.R- project.org/.
Rogers, T. T., & McClelland, J. L. (2005). A Parallel Distributed Processing Approach to Semantic Cognition: Applications to Conceptual Development. In *Building object categories in developmental time* (pp. 335–387). Lawrence Erlbaum Associates Publishers.
Schriefers, H., Meyer, A. S., & Levelt, W. J. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language, 29*, 86–102. https://doi.org/10.1016/0749-596X(90)90011-N
Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review, 81*, 214–241. https://doi.org/10.1037/h0036351
Solomon, S. H., Hindy, N. C., Altmann, G. T. M., & Thompson-Schill, S. L. (2015). Competition between mutually exclusive object states in event comprehension. *Journal of Cognitive Neuroscience, 27*(12), 2324–2338. https://doi.org/10.1162/jocn_a_00866
Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science, 12*(2), 153–156. https://doi.org/10.1111/1467-9280.00326
Stoet, G. (2010). PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods, 42*(4), 1096–1104. https://doi.org/10.3758/BRM.42.4.1096
Stoet, G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-Ttime experiments. *Teaching of Psychology, 44*(1), 24–31. https://doi.org/10.1177/0098628316677643
Symes, E., Ellis, R., & Tucker, M. (2007). Visual object affordances: Object orientation. *Acta Psychologica, 124*(2), 238–255. https://doi.org/10.1016/j.actpsy.2006.03.005
Tabor, W., Juliano, C., & Tanenhaus, M. K. (1997). Parsing in a dynamical system: An attractor-based account of the interaction of lexical and structural constraints in sentence processing. *Language and Cognitive Processes, 12*, 211–271. https://doi.org/10.1080/016909697386853
Trabasso, T., & Suh, S. (1993). Understanding text: Achieving explanatory coherence through on-line inferences and mental operations in working memory. *Discourse Processes, 16*(1–2), 3–34. https://doi.org/10.1080/01638539309544827
van Dijk, T. A., & Kintsch, W. (1983). *Strategies in discourse comprehension*. New York: Academic Press.
Vigliocco, G., Vinson, D. P., Lewis, W., & Garrett, M. F. (2004). Representing the meanings of object and action words: The featural and unitary semantic space hypothesis. *Cognitive Psychology, 48*(4), 422–488. https://doi.org/10.1016/j.cogpsych.2003.09.001

Yaxley, R. H., & Zwaan, R. A. (2007). Simulating visibility during language comprehension. *Cognition, 105*, 229–236. https://doi.org/10.1016/j.cognition.2006.09.003

Yee, E., & Sedivy, J. C. (2006). Eye movements to pictures reveal transient semantic activation during spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 1–14. https://doi.org/10.1037/0278-7393.32.1.1

Yee, E., & Thompson-Schill, S. L. (2016). Putting concepts into context. *Psychonomic Bulletin & Review, 23*(4), 1015–1027. https://doi.org/10.3758/s13423-015-0948-7

Zwaan, R. A. (1996). Processing narrative time shifts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1196–1207. https://doi.org/10.1037/0278-7393.22.5.1196

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science, 292–297*. https://doi.org/10.1111/j.1467-9280.1995.tb00513.x

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language and memory. *Psychological Bulletin, 123*, 162–185. https://doi.org/10.1037/0033-2909.123.2.162

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language Comprehenders Mentally Represent the Shapes of Objects. *Psychological Science, 13*(2), 168–171. https://doi.org/10.1111/1467-9280.00430